



**Faculté de Médecine,
Pharmacie**

Conférence des Doyens
des facultés de Médecine



SANTÉ
FORMATION
RECHERCHE

**UNIVERSITE
DE ROUEN**

Le Test de Concordance de Script Pourquoi ? Comment ?

Louis SIBERT - Rouen -

Séminaire du groupe pédagogique de la conférence des Doyens- 07/10/2020-

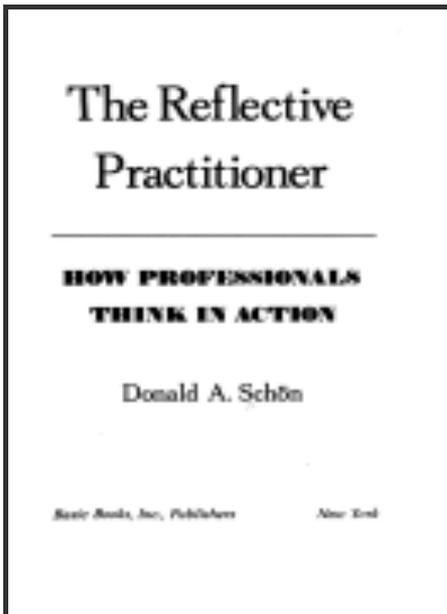
De quoi parle-t-on ?

- **Mesurer un aspect particulier du raisonnement clinique :**
 - l'aptitude à interpréter des informations médicales dans un **contexte d'incertitude**.
- **Originalités :**
 - **Comparaison des choix des étudiants à ceux d'un panel d'experts** : plus leurs choix se rapprochent de ceux des experts, plus le raisonnement des étudiants est considéré comme performant
 - Intègre **la notion d'incertitude** : recommandations non applicables à un cas particulier, etc...
- **Caractéristiques :**
 - Technique d'évaluation **standardisée**
 - Format **docimologiquement validé**
 - Repose sur des **bases conceptuelles solides**

Doit-on évaluer la gestion de l'incertitude en sciences de la santé ?

- Environ 80 p. cent des décisions en médecine sont prises en contexte d'incertitude
- Une des différences majeures entre experts et juniors est l'aptitude des experts à prendre une décision **en contexte d'incertitude**, quel que soit le domaine
- Corrélation entre le manque d'entraînement à la gestion de l'incertitude et
 - augmentation des prescriptions d'examens complémentaires
 - Augmentation des prescriptions hors recommandations professionnelles

- *Cooke S. Acad Med 2017*

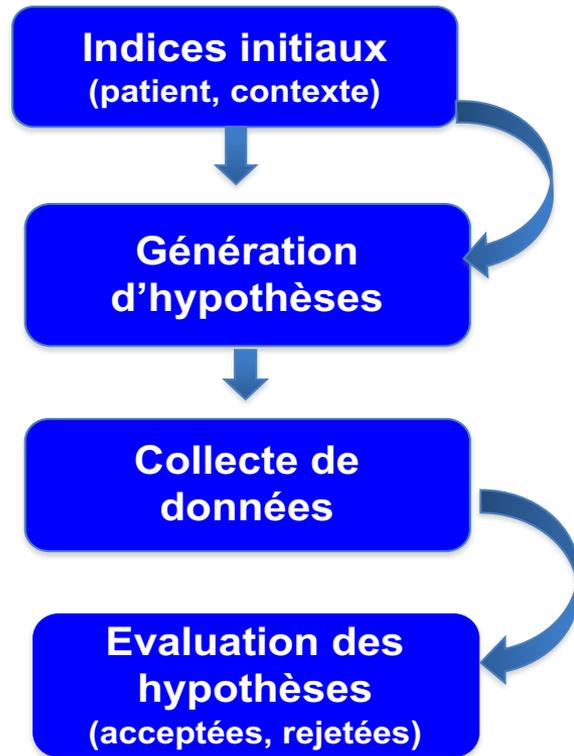


TCS : structure élémentaire

**Vous vous apprêtez à partir en voiture le matin du 11 février 2013.
Il neige. Votre voiture est garée dans la rue. Elle ne démarre pas**

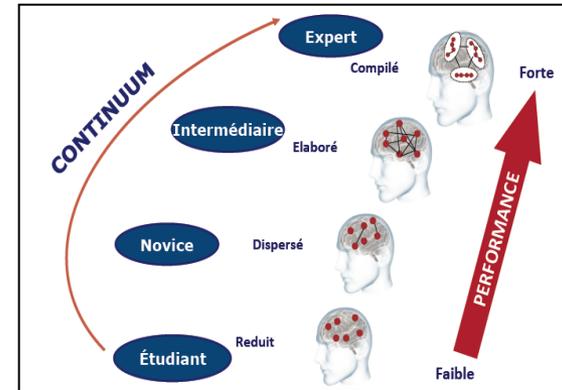
Si vous pensiez :	Et qu'alors vous constatez :	L'impact de cette nouvelle information sur votre hypothèse est :
Que votre réservoir est vide	Vous avez roulé la veille avec le témoin du réservoir allumé	-2 -1 0 +1 +2
Que votre batterie est morte	Vous avez acheté votre voiture neuve il y a 4 ans	-2 -1 0 +1 +2
Que votre alternateur ne fonctionne pas	Rien ne s'allume au démarrage	-2 -1 0 +1 +2
-2: Totalemnt négatif 0 : Ni plus ni moins négatif		+1 : Positif +2 : Très positif

Raisonnement hypothético-déductif



- Schmidt HG. Acad Med 1990
- Schmidt HG. Med Educ 2007

Théorie des scripts



- Organisation des connaissances en réseaux spécifiquement adaptés pour les tâches cliniques
- Dès le début de l'exposition clinique
- Multiples micro-jugements au cours du processus de prise de décision, peuvent faire l'objet de mesure et comparaison

- Barrows HS. Med Educ 1987
- Charlin B. Acad Med 2000
- Schmidt HG. Med Educ 2007

- Situation d'incertitude
- Pas de bonne réponse
- Dans quelles mesure les décisions prises par les candidats **CONCORDENT** avec celles des membres du panel
- Points fonction du nombre de membres du panel qui ont fourni la même réponse
 - Score maximum : réponse du plus grand nombre d'experts (**réponse modale, majoritaire**)
 - Autres réponses d'experts : différences d'interprétations, **cliniquement pertinentes**

Scores et notation

www.cme.umontreal.ca/tcs

Coefficient de fidélité

Fidélité		intervalle 95%		faible corrélation 0,05		bonne corrélation 0,15		critères		variance faible 0,50		variance élevée 1,00	
item variance	12,4												
test variance	88,3												
nombre d'items	90												
Cronbach alpha	0,87												
nouveau nombre d'items		132											
Nouvel alpha avec N questions similaires		0,91											
					faible 6	moyenne 9	bonne 75					nombres	29 17

Item statistics	mean	sd	variance	correlation item/total	Panel responses											
					faible	moyenne	bonne	1	2	3	4	5	mode	variance	low variance	high variance
q1	0,429	0,356	0,126	0,36			✓	8	23	10	2	2	2	0,93		
q2	0,299	0,410	0,168	0,46			✓	39	5	1	0	0	1	0,18	X	
q3	0,344	0,310	0,096	0,18			✓	6	14	22	0	3	3	0,93		
q4	0,477	0,420	0,176	0,20			✓	1	4	29	5	6	3	0,78		
q5	0,725	0,435	0,190	0,37			✓	0	2	40	1	2	3	0,25	X	
q6	0,533	0,478	0,228	0,40			✓	0	2	39	2	2	3	0,26	X	
q7	0,501	0,483	0,234	0,36			✓	1	2	39	1	2	3	0,34	X	
q8	0,577	0,411	0,169	0,39			✓	7	24	13	1	0	2	0,51		
q9	0,598	0,433	0,187	0,57			✓	4	22	18	1	0	2	0,46	X	
q10	0,736	0,413	0,170	0,38			✓	1	34	9	1	0	2	0,27	X	
q11	0,719	0,363	0,132	0,26			✓	0	0	3	31	11	4	0,29	X	
q12	0,707	0,413	0,170	0,15			✓	0	0	1	21	23	5	0,30	X	
q13	0,588	0,433	0,187	0,24			✓	1	3	7	28	6	4	0,72		
q14	0,611	0,359	0,129	0,24			✓	2	7	3	8	5	1	2,25		X

les nombres en vert peuvent être modifiés

Niveau de discrimination de chaque question

Réponses du panel d'experts

TCS : principes de rédaction

*Fournier JP, BMC Med inf & Decision Making 2008 ;
Dory V, Med Educ 2012*

1) Préparation : 1 à 2 rédacteurs

- **Choix des situations :**
 - **Discipline** : urologie, **Domaine** : diagnostic, **Cible** DFSAM1
 - **Situation** : augmentation progressive de volume d'une bourse chez l'adulte
- **Rédaction des vignettes cliniques :**
 - Un patient de 24 ans vient vous consulter car il a constaté une augmentation de volume de sa bourse gauche depuis 1 mois, sensible depuis 48 heures
- **Choix des hypothèses**
 - Tumeur du testicule, Hydrocèle, Varicocèle
- **Choix des informations supplémentaires**
 - **Tumeur** : induration, microlithiases à l'échographie, cryptorchidie, ...
 - **Hydrocèle** : transillumination, antécédents d'infection urinaire, orifices herniaires libres, ...
 - **Varicocèle** : douleur, hypotrophie testiculaire, reflux en échographie-doppler

TCS : principes de rédaction

*Fournier JP, BMC Med inf & Decision Making 2008 ;
Dory V, Med Educ 2012*

Un patient de 24 ans vient vous consulter car il a constaté une augmentation de volume de sa bourse gauche depuis 1 mois, sensible depuis 48 heures

Si vous pensiez à :	Et qu'alors vous trouvez :	L'impact de cette nouvelle information sur votre hypothèse est :
Tumeur du testicule	Antécédent de cure chirurgicale d'anomalie de descente testiculaire dans l'enfance	-2 -1 0 +1 +2
Hydrocele	Les orifices herniaires sont libres	-2 -1 0 +1 +2
Varicocele	Douleurs scrotales chroniques avec sillon épидидymo testiculaire bien perçu à l'examen clinique	-2 -1 0 +1 +2
-2: Totalement négatif		0 : Ni plus, ni moins positif
-1 : Négatif		+1 : Positif
		+2: Très positif

TCS : principes de rédaction

*Fournier JP, BMC Med inf & Decision Making 2008 ;
Dory V, Med Educ 2012*

2) Validation / vérification : 1 ou 2 “experts” de la méthode

- **Situations** : pertinentes à la pratique, adaptées au programme
- **Vignettes cliniques** : incertitude, sexe, age, lieu, motif de recours au soin, +/- clinique, +/- imagerie, biologie..
- **Hypothèses** : celles que générerait un clinicien expérimenté dans le contexte défini par la vignette
- **Informations supplémentaires** : celles qu’un clinicien expérimenté chercherait dans le contexte défini par la vignette, compte-tenu des hypothèses

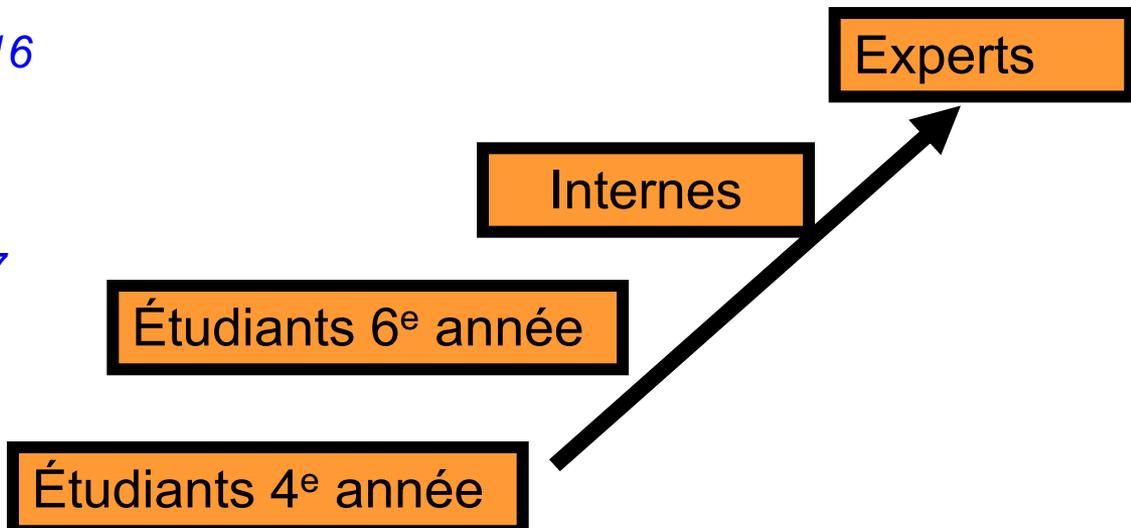
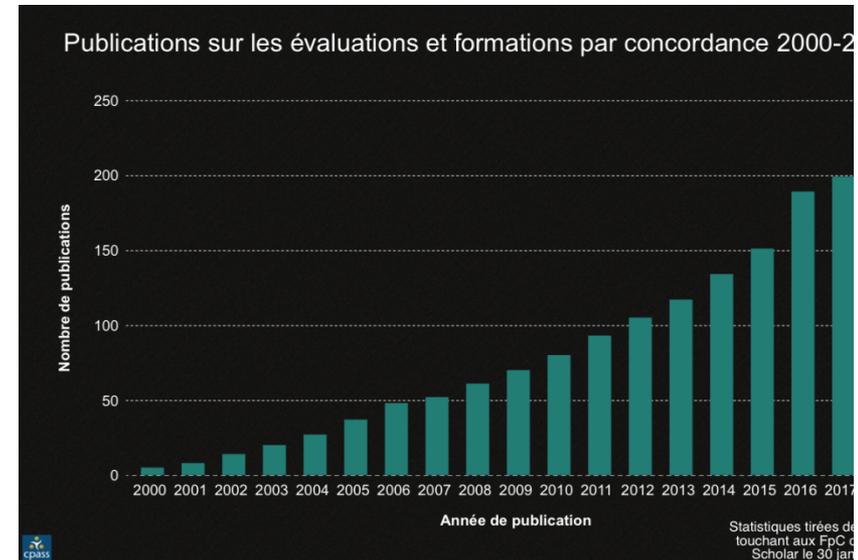
3) Etablissement des scores : Panel d’experts

- Passent le test individuellement

Qualités psychométriques du TCS

- Sibert L. *Eur Urol* 2002; *Med Teacher* 2002
- Caire F, *Neurochirurgie* 2004
- Meterissian S. *Am J Surg* 2007
- Gibot S. *Pédagogie Médicale* 2008
- Carriere B. *Ann Emerg Med* 2009
- Sibert L. *Med Teacher* 2009
- Burstejn A. *J Eur Acad Dermatol Venerol* 2011
- Duggan P, *BMC Med Educ* 2012
- Nouh T. *Am J Surg* 2012
- Piorezan KD. *J Am Geriatr Soc* 2012
- Dory V. *Fam Med* 2014
- Compère V. *J Clin Anesthesia* 2016
- Goos M, *BMC Surgery* 2016
- Kazour F, *Acad Psychiatry* 2016
- Sibert L, *Rev Mal Resp* 2016
- Chime NO, *Ped Emerg Care* 2017
- Pau A, *J Acad Ethics* 2019

Coefficient alpha :
0.75 à 0.86



Comparaison aux autres formats

- Comparaison TCS et QCM à contexte riche en médecine d'urgence
- 60 QCM, 30 TCS, 20 internes, 16 DCEM4, 9 urgentistes
- *Fournier JP, Pédagogie Médicale 2006*

Figure 1 :

Scores (sur 100) obtenus par les différentes populations au test de concordance des scripts

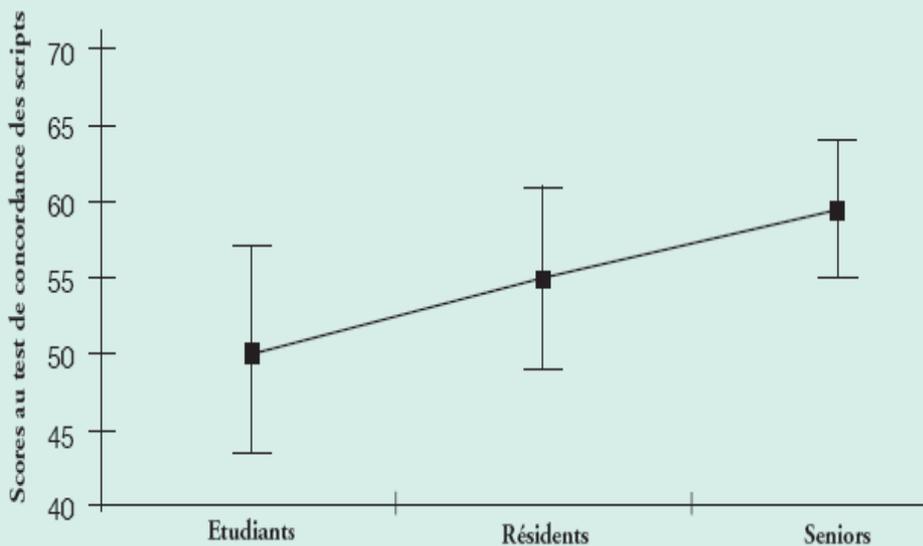
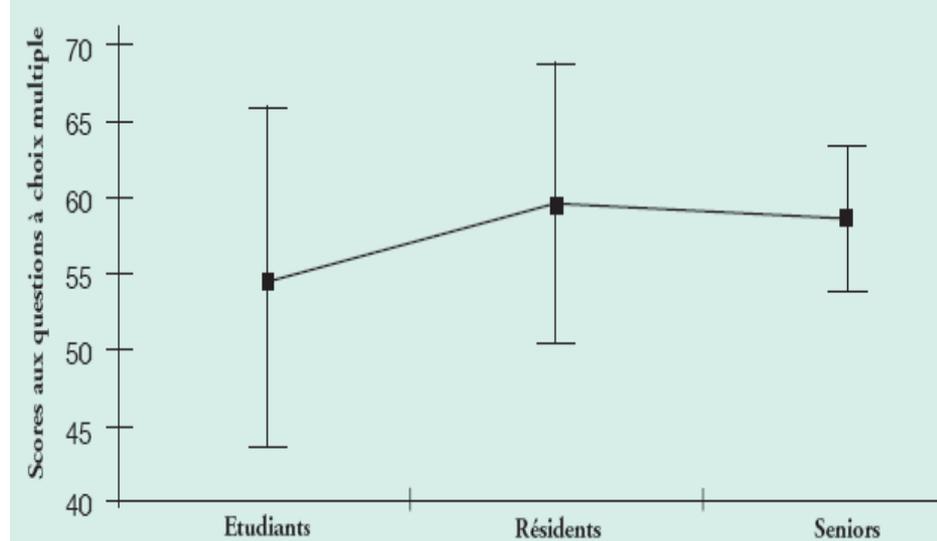


Figure 2 :

Scores (sur 100) obtenus par les différentes populations aux questions à choix multiple



Seul le TCS identifie le niveau d'expérience

Comparaison aux autres formats

Teaching and Learning in Medicine, 24(3), 187–193
Copyright © 2012, Taylor & Francis Group, LLC
ISSN: 1040-1334 print / 1532-8015 online
DOI: 10.1080/10401334.2012.692239

APPLIED RESEARCH

Comparing a Script Concordance Examination to a Multiple-Choice Examination on a Core Internal Medicine Clerkship

William Kelly, Steven Durning, and Gerald Denton

Department of Medicine, Uniformed Services University of Health Sciences, Bethesda, Maryland, USA

- 163 étudiants 3e année
- Médecine interne
- 20 QCM, 20 TCS

Corrélation scores avec performance clinique en stage

	NBME	MCQ	SC	Clinical Performance
NBME		0.352 <i>p</i> < .001	0.352 <i>p</i> < .001	0.484 <i>p</i> < .001
MCQ			0.360 <i>p</i> < .001	0.111 <i>p</i> = .159
SC				0.221 <i>p</i> = .005

Clinical Reasoning Assessment Methods: A Scoping Review and Practical Guidance.

Daniel M. Acad Med. 2019;94:902-912



Assessment method: Definition	Clinical reasoning component							IG : information Gathering
	IG	HG	PR	DD	LD	DJ	MT	
Non-workplace-based assessments								HG : Hypothesis Generation
Clinical or comprehensive integrative puzzles: An extended matching crossword puzzle designed to assess a learner's ability to relate clinical vignettes to specific diagnoses and diagnostic or therapeutic interventions.	0.4	0.3	0.6	1.1	1.9	0.4	1.3	PR : Pb repres.
Concept maps: A schematic method for learners to organize and represent their knowledge and knowledge structures through a graphical illustration of the complex processes and relationships between concepts within a subject domain.	0.4	0.4	1.2	1.0	0.4	0.8	0.9	DD : Differential Dg
Extended matching questions: A written exam format consisting of a lead-in question (clinical vignette) followed by multiple answer options in a list where more answer options are given than in multiple-choice questions (i.e., > 5).	0.2	0.3	0.2	0.8	1.7	0.3	1.3	LD : Leading dg
Key feature examinations: Problems typically consist of a clinical vignette followed by 2–3 questions that assess the critical elements ("key features") or challenging decisions that clinicians must make.	0.9	0.5	0.4	1.5	1.4	0.6	1.4	DJ : Dg justifications
Multiple-choice questions: A clinical vignette is followed by up to 5 alternatives. Questions may take the following formats: single best alternative, matching, true or false, and combinations of alternatives.	0.9	0.3	0.0	0.6	1.9	0.0	1.8	MT : Management treatment
Modified essay questions: A method wherein serial information about a clinical case is presented chronologically. After each item, the learner must document a decision. The student cannot preview subsequent items until a decision is made.	1.3	1.2	1.0	1.6	1.7	1.3	1.7	
Oral examinations: A verbal examination conducted by one or more faculty members through unscripted or semiscripted questions that assess clinical reasoning and decision-making abilities, as well as professional values.	1.3	1.3	1.1	1.8	1.8	1.9	1.9	
Patient management problems: A clinical scenario is presented in real-life settings with specific resources available for diagnosis or management. The learner chooses among multiple alternatives. The results of actions (e.g., labs, images) are provided.	1.6	1.0	0.3	1.4	1.9	0.6	1.7	
Script concordance tests: Clinical scenarios with uncertainty are followed by a series of questions (e.g., if you are thinking X and you find Y, the answer becomes more likely, less likely, or no change). Responses are compared with those of experts.	0.4	0.8	0.6	0.8	1.3	0.9	1.1	
Short- or long-answer (essay) questions: A clinical vignette is followed by one or more questions. Learners provide free-text responses that range in length from a few words to several sentences.	0.8	1.2	1.2	1.8	1.7	1.8	1.7	

TCS au cours du 2^e cycle

Mathieu et al *BMC Medical Education* 2013, **13**:166
<http://www.biomedcentral.com/1472-6920/13/166>



RESEARCH ARTICLE

Open Access

Construction and utilization of a script concordance test as an assessment tool for dcem3 (5th year) medical students in rheumatology

Sylvain Mathieu^{1,3*}, Marion Couderc¹, Baptiste Glace¹, Anne Tournadre¹, Sandrine Malochet-Guinamand¹, Bruno Pereira², Jean-Jacques Dubost¹ and Martin Soubrier¹

Assessment of Undergraduate Clinical Reasoning in Geriatric Medicine: Application of a Script Concordance Test

Ronaldo D. Piovezan, MD, MSc,^{*}† Osvladir Custódio, MD, MSc,^{*} Maysa S. Cendoroglo, MD, PhD,^{*} Nildo A. Batista, MD, PhD,[†] Stuart Lubarsky, MD, MHPE,[‡] and Bernard Charlin, MD, PhD[§]

2011; 33: 472-477



Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students

ALOYSIUS J. HUMBERT¹, MARY T. JOHNSON², EDWARD MIECH³, FRED FRIEDBERG⁴, JANICE A. GRACKIN⁵ & PEGGY A. SEIDMAN⁶

¹Indiana University School of Medicine, USA, ²Florida State University College of Medicine, USA, ³Roudebush Veteran Affairs Medical Center, USA, ⁴Department of Psychiatry, Stony Brook University Medical Center, USA, ⁵Nassau Community College, USA, ⁶Department of Anesthesiology and Pediatrics, Stony Brook University Medical Center, USA

Measuring Gains in the Clinical Reasoning of Medical Students: Longitudinal Results From a School-Wide Script Concordance Test

Aloysius J. Humbert, MD, and Edward J. Miech, EdD

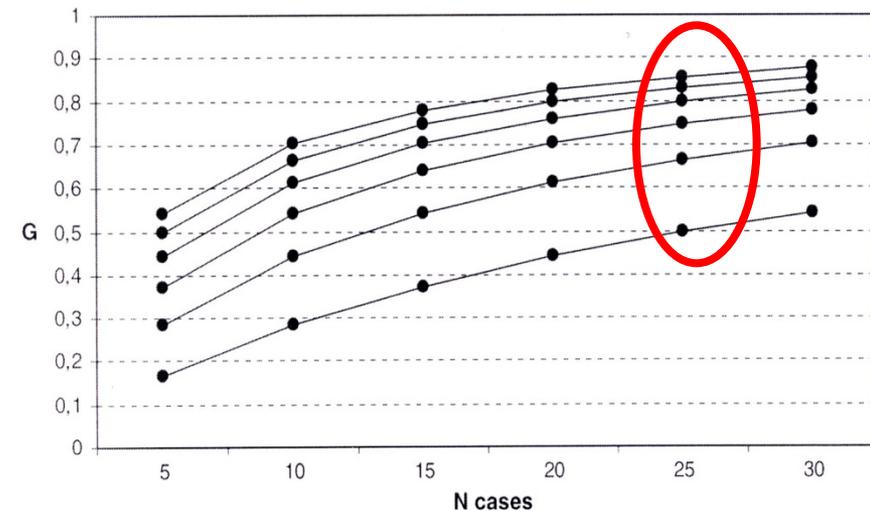
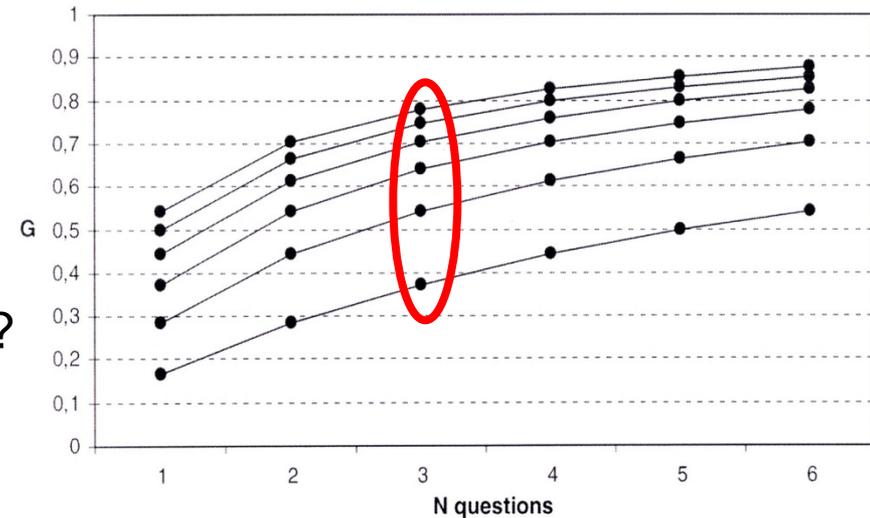
Combien de cas et de questions ?

- Nombreux cas avec 1 question ?
- Moins de cas avec plusieurs questions ?
- Combien de cas pour une bonne fiabilité ?

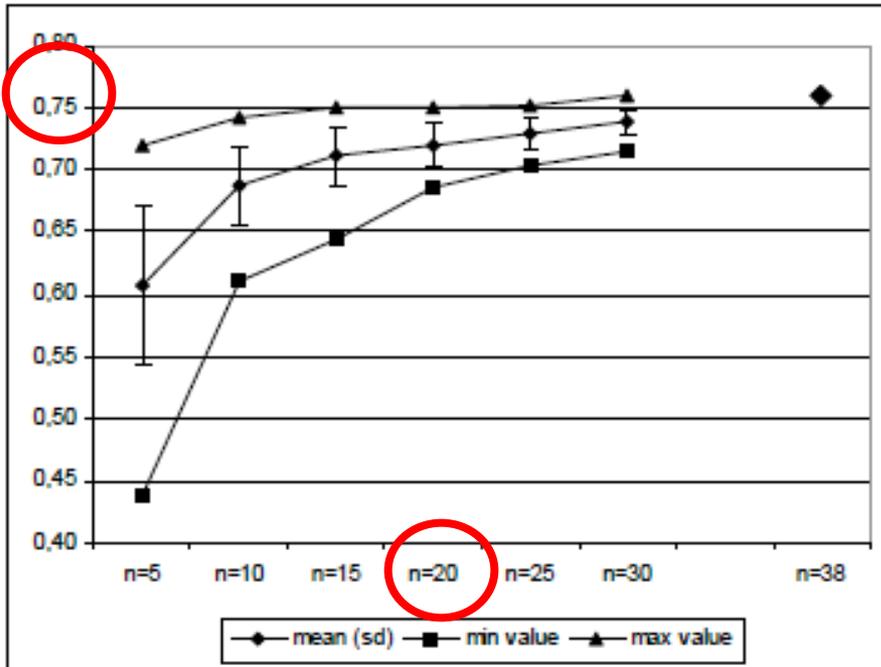
- **20 – 25 cas par examens**
- **3 questions par cas**
- **60 mn d'examen**
- **Score alpha > 0,70**

Gagnon R.

Adv Health Sci Educ Theory Pract. 2008

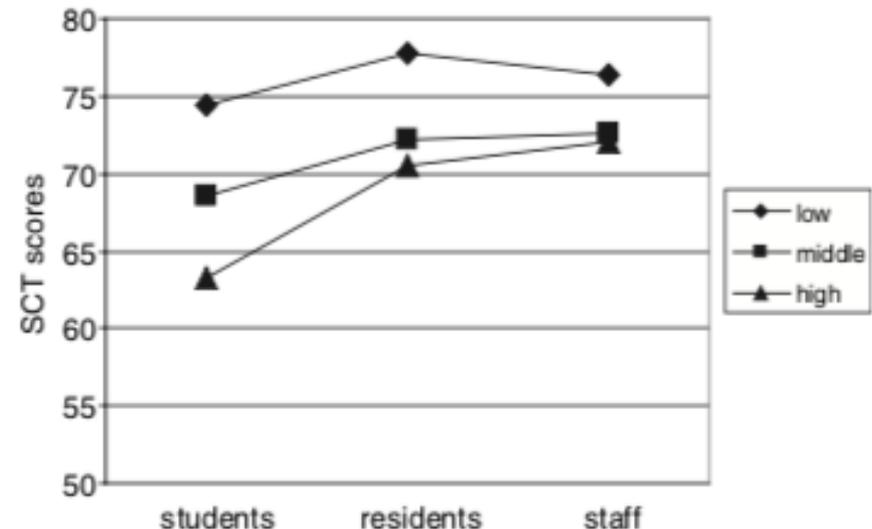


Panel : combien de membres ?, quelle variabilité des réponses ?



- *Gagnon R. Med Educ 2005*

Variabilité correcte : Réponses attendues (et observées) réparties autour de la valeur modale (1 à 2 graduations de l'échelle de Likert)



- *Charlin B. Medical Education 2006*
- *Gagnon R. Adv in Health Sci Educ 2011*

Impact des réponses au hasard ou du choix systématique de la valeur 0

Effects of Gaming and Guessing on Script Concordance Test Score. Lubarski S, AMEE 2015

- Réponse au hasard : - 2 DS / score moyen
- Réponse “0” systématique : +1 DS / score moyen

- Pas d’impact positif des réponses au hasard
- Limiter les questions à réponse modale attendue “0” à > 10% des questions

Impact de la composition du panel d'experts sur les scores en cas de TCS multi disciplinaire ?

	Panels	Candidats	TCS	Score α
Humbert <i>Med Teach 2011</i>	30 internistes	481 étudiants	75 items	0,73
Duggan <i>BMC Med Educ 2012</i>	7 disciplines 10 à 16 experts/ disciplines	158 étudiants	51 cas 158 items	0,78
Nouh <i>Am J Surg 2012</i>	22 chir. générale	202 résidents	53 cas 153 items	0,85
Petrucci <i>Am J Surg 2012</i>	5 disciplines chirurgicales 5 experts / disc.	227 résidents	53 cas 153 items	0,81

Apport de SIDES : Choix des experts

Ajouter des experts au panel

Rouen

Caen

Caen

Lorraine Médecine

Nice

0 - Plateforme Nationale des Disciplines

Amiens

Angers

Antilles et Guyane

Besançon

Bordeaux Médecine

Brest

Caen

Clermont-Ferrand

Dijon

Grenoble Médecine

La réunion

Lille

Lille Catho

Limoges

Lorraine Médecine

Lyon 3ème cycle

Lyon Est

Lyon Sud

Marseille Médecine

Montpellier

Nantes

Nice

Paris-Ile de France Ouest

Paris V

Paris VI

Paris VII

Paris XI

Paris XII

Paris XIII

Poitiers

Reims

Rennes

Saint-Etienne

Strasbourg

Toulouse Purpan

Toulouse Rangueil

Tours

Catégorie du groupe

Groupe

Nom

Prénom

Spécialités

Filter

Réinitialiser

- Choix :
 - Local
 - Régional
 - national

Liste des experts disponibles

Suivi du panel d'experts

Ajouter des experts

Liste des experts ajoutés

manon Allaire
 Terminé

Rodolphe Anty
 Terminé

Arthur Belle
 Terminé

Clémence Canivet
 Terminé

Thong Dao
 Terminé

Frédéric Di Fiore
 Pas encore commencé

0

Relancer tous les experts qui n'ont pas encore répondu

Retirer les experts qui n'ont pas été contactés

Envoyer un mail aux experts qui n'en ont jamais reçu

Nombre d'experts dans le panel : 12

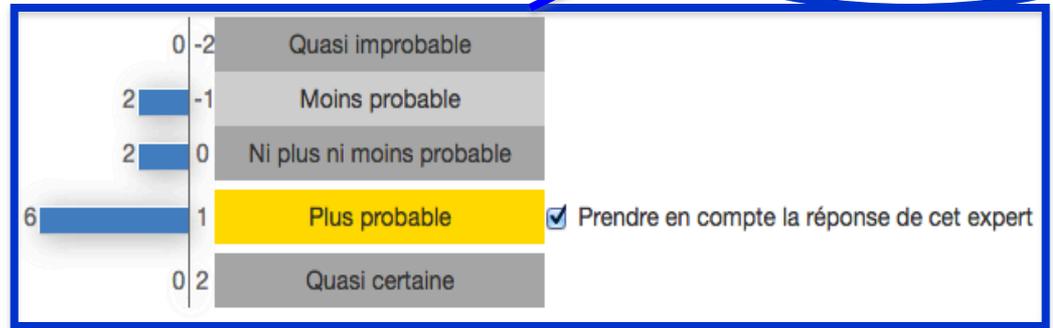
Nombre d'experts du panel ayant répondu : 11

theia.isd.question_list.question_num

Si vous pensiez à ...
 Une sigmoïdite diverticulaire

Et qu'alors vous ...
 avez trouvé une boîte d'ibuprofène dans son
 sac à main

Votre hypothèse initiale devient...



- Suivi des experts
- Relance
- Elimination
- Nouveaux experts

Validation des réponses individuelles

TCS blanc national ?

- **Multi-facultaires** : Caen, Nancy, Nice, Rouen, Paris 5, Paris 7, Paris 6, Paris 13, Paris 11, UVSQ, Lyon est, Lyon sud, Nantes, Montpellier, ...
- **Etudiants de DFASM 2**
- **Matières communes** : Cardio, Pneumo, Gastro, Neuro, Psychiatrie (à confirmer) ; 25 vignettes, 75 questions, Etudiants volontaires
- **Experts** :
 - spécialistes vs experts polyvalents
 - « Jeunes experts » vs « vieux » experts
- **Questions de recherche** :
 - Reproductibilité des scores : fiabilité interne, généralisabilité
 - Stabilité des scores en test-retest à grande échelle
 - Comparaison des scores pour la même épreuve avec jury polyvalent vs jury par spécialité
 - Variation des panels
 - Etc...

BMC Medical Informatics and Decision Making



Correspondence

Script Concordance Tests: Guidelines for Construction

Jean Paul Fournier¹, Anne Demeester² and Bernard Charlin^{*3}

Open Access

2013, 1–10, Early Online



AMEE GUIDE

Script concordance testing: From theory to practice: AMEE Guide No. 75

STUART LUBARSKY¹, VALÉRIE DORY², PAUL DUGGAN³, ROBERT GAGNON⁴ & BERNARD CHARLIN⁴

¹McGill University, Canada, ²Université catholique de Louvain, Belgium, ³University of Adelaide, Australia, ⁴University of Montreal, Canada

Louis SIBERT, Jean-Paul FOURNIER

ECNi – Épreuve TCS Test de concordance de script

Le guide méthodologique

CERTIFICATION SIDES



Système Informatique Distribué
d'Évaluation en Santé

Le Test de Concordance de Script (TCS):
bases conceptuelles, principes d'élaboration
et d'administration

Louis Sibert¹, Jean Paul Fournier², Bernard Charlin³

Le test de concordance de script en 20 questions

Twenty questions on script concordance tests

Didier GIET^{1,3}, Valérie MASSART¹, Robert GAGNON² et Bernard CHARLIN²

Contrôle qualité *a priori*

Vignette	<ul style="list-style-type: none"> • Décrit une situation problématique, même pour les experts • Décrit une situation clinique adaptée au niveau de compétences des étudiants • Question et contexte sont explicites • Présentation clinique typique • Vignette correctement rédigée 	<input type="checkbox"/> Oui <input type="checkbox"/> non <input type="checkbox"/> Oui <input type="checkbox"/> non
Questions	<ul style="list-style-type: none"> • Questions focalisant sur les points importants • Hypothèses pertinentes • 2 options identiques non rencontrées dans 2 questions consécutives • Information supplémentaire pertinente/hypothèse • Ancrages explicites, en particulier la valeur neutre (0) • Questions construites pour balayer les 5 points de l'échelle de Likert • Questions construites pour obtenir une variabilité moyenne 	<input type="checkbox"/> Oui <input type="checkbox"/> non <input type="checkbox"/> Oui <input type="checkbox"/> non
Panel d'experts	<ul style="list-style-type: none"> • Entre 10 et 20 • Médecins expérimentés dont la présence dans le panel est licite compte tenu du niveau de compétences des étudiants • Les experts passent le test dans les mêmes conditions que les étudiants 	<input type="checkbox"/> Oui <input type="checkbox"/> non <input type="checkbox"/> Oui <input type="checkbox"/> non <input type="checkbox"/> Oui <input type="checkbox"/> non

Caire, Neurochirurgie 2004

Le TCS m'a tuer...



TCS

Processus de
réponses

Etablissement
des scores

Constitution des
panels

Inadaptés au
second cycle